

On the \top -Stein equation $X = AX^\top B + C$

Matthew M. Lin¹

Department of Mathematics, National Chung Cheng University, Chia-Yi 621, Taiwan.

Chun-Yueh Chiang^{2,*}

Center for General Education, National Formosa University, Huwei 632, Taiwan.

Abstract

We consider the \top -Stein equation $X = AX^\top B + C$, where the operator $(\cdot)^\top$ denotes the transpose (\top) of a matrix. In the first part of this paper, we analyze necessary and sufficient conditions for the existence and uniqueness of the solution X . In the second part, a numerical algorithm for solving \top -Stein equation is given under the solvability conditions.

Keywords: Sylvester equation, Stein equation, PQZ decomposition

1. Introduction

The purpose of this article is to analyze the solvability conditions of the \top -Stein equation

$$X = AX^\top B + C, \quad (1)$$

where $A, B, C \in \mathbb{C}^{m \times n}$ are known matrices, $X \in \mathbb{C}^{m \times n}$ is a matrix to be determined. Our interest in the \top -Stein equation originates from the study of completely integrable mechanical systems, that is, the analysis of the \top -Sylvester equation

$$AX + X^\top B = C, \quad (2)$$

where A, B, X are matrices in $\mathbb{R}^{m \times n}$ [1, 2]. By means of the generalized inverses or QZ decomposition [3], the solvability conditions of (2) are studies in [1, 2, 4]. Suppose that (A, B^\top) is regular, that is, $aA + bB^\top$ is invertible for some scalars a and b . The \top -Sylvester equation (2) can be written as

$$(aA + bB^\top)X + X^\top(aB + bA^\top) = aC + bC^\top. \quad (3)$$

*Corresponding author

Email addresses: mllin@math.ccu.edu.tw (Matthew M. Lin), chiang@nfu.edu.tw (Chun-Yueh Chiang)

¹The first author was supported by the National Science Council of Taiwan under grant 101-2115-M-194-007-MY3.

²The second author was supported by the National Science Council of Taiwan under grant NSC100-2115-M-150-001.

Premultiplying both sides of (3) by $(aA + bB^\top)^{-1}$, we have

$$X + UX^\top V = W, \quad (4)$$

where $U = (aA + bB^\top)^{-1}$, $V = aB + bA^\top$ and $W = (aA + bB^\top)^{-1}(aC + bC^\top)$. This is of the form (1). In other words, numerical approaches for solving (2) can be obtained by transforming (2) into the form of (1), and then applying some iterative methods to (1) for the solution [4, 5, 6]. With this in mind, in this note we are interested in the study of \top -Stein matrix equation (1).

Observe that if X is a solution of (1), then X also satisfies the following Stein equation

$$X = AB^\top XA^\top B + C + AC^\top B. \quad (5)$$

It follows that once there exists a unique solution to (5), we can usually solve (1) via the conventional Smith-type iterative methods given in [7, 8] to (5). However, even if (1) is uniquely solvable, it does not imply (5) is uniquely solvable. For example, if $b = 1$ and $a = -1$, the scalar equation $x = -x + c$ has a unique solution $x = \frac{c}{2}$, but no more information can be obtained from the Stein equation $x = x + c - c$. In our work, we formulate the necessary and sufficient conditions for the existence of the solution of (1) directly by means of the spectrum analysis in Section 2. We study the capability of the Smith-type iterative algorithms for solving (1) in Section 3 and concluding remarks are given in Section 4.

2. Solvability conditions of the Matrix Equation $X = AX^\top B + C$

In order to formalize our discussion, let the notations $A \otimes B$ be the Kronecker product of matrices A and B , I_n be the $n \times n$ identity matrix, and $\sigma(A)$ be the set of all eigenvalues of A . With the Kronecker product, (1) can be written as the enlarged linear system

$$(I_{mn} - (B^\top \otimes A)\mathcal{P})\text{vec}(X) = \text{vec}(C), \quad (6)$$

where $\text{vec}(X)$ stacks the columns of X into a column vector and \mathcal{P} is the Kronecker permutation matrix [9] which maps $\text{vec}(X)$ into $\text{vec}(X^\top)$, i.e., $\mathcal{P} = \sum_{1 \leq i, j \leq mn} e_j e_i^\top \otimes e_i e_j^\top$, where e_i denotes the i -th column of the $mn \times mn$ identity matrix I_{mn} . Due to the specific structure of \mathcal{P} , it has been shown in [10, Corollary 4.3.10] that

$$\mathcal{P}^\top (B^\top \otimes A) \mathcal{P} = A \otimes B^\top.$$

It then follows that

$$((B^\top \otimes A)\mathcal{P})^2 = (B^\top \otimes A)\mathcal{P}\mathcal{P}^\top (A \otimes B^\top) = B^\top A \otimes AB^\top, \quad (7)$$

since $\mathcal{P}^2 = I_{mn}$ and $\mathcal{P} = \mathcal{P}^\top$. Note that $\sigma(A^\top B) = \sigma(B^\top A) = \sigma(AB^\top)$. By (7) and the property of the Kronecker product [11, Theorem 4.8], we know that $\sigma(((B^\top \otimes A)\mathcal{P})^2) = \{\lambda_i \lambda_j | \lambda_i, \lambda_j \in \sigma(A^\top B) = \{\lambda_1, \dots, \lambda_n\}, 1 \leq i, j \leq n\}$.

That is, the eigenvalues of $(B^\top \otimes A)\mathcal{P}$ is related to the square roots of the eigenvalues of $\sigma(A^\top B)$, but from (7), no more information can be used to decide the positivity or non-negativity of the eigenvalues of $(B^\top \otimes A)\mathcal{P}$. A question immediately arises as to whether it is possible to obtain the explicit expression of the eigenvalues of $(B^\top \otimes A)\mathcal{P}$, provided the eigenvalues of $A^\top B$ are given. In the following two Lemmas, we first show that the generalized inverse of an upper triangular matrix is still upper triangular and then apply it to discuss the eigenvalues of $(B^\top \otimes A)\mathcal{P}$.

Lemma 2.1. *Suppose that $A \in \mathbb{C}^{m \times n}$ is an upper triangular matrix having full rank. Then the generalized inverse of A is an upper triangular matrix.*

PROOF. Let $A = \begin{bmatrix} \hat{A}^\top & 0 \end{bmatrix}^\top$, if $m \geq n$, or $A = \begin{bmatrix} \hat{A} & * \end{bmatrix}$, if $m \leq n$. Here, \hat{A} is an upper triangular matrix with $\text{rank}(\hat{A}) = \text{rank}(A)$. Then, the generalized inverse A^\dagger of A is an $n \times m$ matrix satisfies $A^\dagger = \begin{bmatrix} \hat{A}^{-1} & 0 \end{bmatrix}$, if $m \geq n$, or $A^\dagger = \begin{bmatrix} \hat{A}^{-T} & 0 \end{bmatrix}^\top$ if $m \leq n$. This concludes that A^\dagger is upper triangular.

Lemma 2.2. *Let A and B be two matrices in $\mathbb{C}^{m \times n}$. Then*

1. *There exist unitary matrices $P \in \mathbb{C}^{m \times m}$ and $Q \in \mathbb{C}^{n \times n}$ such that $U_A := PAQ$ and $U_B := Q^H B^\top P^H$ are two upper triangular matrices.*
2. *$(B^\top \otimes A)\mathcal{P} = (Q \otimes P^H)(U_A \otimes U_B)\mathcal{P}(Q^H \otimes P)$*
3. *$\sigma((B^\top \otimes A)\mathcal{P}) = \{\lambda_i, \pm\sqrt{\lambda_i \lambda_j} | \lambda_i, \lambda_j \in \sigma(A^\top B) = \{\lambda_1, \dots, \lambda_n\}, 1 \leq i < j \leq n\}$.*

Here, \sqrt{z} denotes the principal square root of a complex number z .

PROOF. Assume first that B is full rank and $PAB^\top P^H = T_1$ is the upper schur form of AB^\top . This defines the unitary matrix P and one can consider the QR and QL decomposition of matrices $B^\top P^H$ and $(PA)^H$, respectively, such that

$$\begin{aligned} Q^H B^\top P^H &= T_2, & \text{for } m \geq n, \\ \hat{Q}^H (PA)^H &= \hat{T}_2^\top, & \text{for } m < n, \end{aligned}$$

where Q and \hat{Q} are unitary, and T_2 and \hat{T}_2 are upper triangular. By Lemma 2.1, there exist two generalized inverses T_2^\dagger and \hat{T}_2^\dagger of matrices T_2 and \hat{T}_2 , respectively, such that $T_2 T_2^\dagger = I_n$ and $\hat{T}_2^\dagger \hat{T}_2 = I_m$. In turn, one shall have the following upper triangular matrices:

$$\begin{aligned} U_A &:= PAQ = T_1 T_2^\dagger, & U_B &:= Q^H B^\top P^H = T_2, & \text{for } m \geq n, \\ U_A &:= PA\hat{Q} = \hat{T}_2, & U_B &:= \hat{Q}^H B^\top P^H = \hat{T}_2^\dagger T_1, & \text{for } m < n. \end{aligned}$$

If B is not full rank, one then considers the SVD decomposition of B :

$$B = U \text{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0) V^\top,$$

where $U \in \mathbb{C}^{m \times s}$ and $V \in \mathbb{C}^{s \times n}$ are unitary, $s = \min(m, n)$, and $r = \text{rank}(B)$. This then defines a sequence of nonsingular matrices

$$B_k := B + \frac{1}{k} \sum_{i=r+1}^s U_i V_i^\top,$$

for any positive integer k . It follows from the discussion above that for any positive integer k , there exist unitary matrices $P_k \in \mathbb{C}^{m \times m}$ and $Q_k \in \mathbb{C}^{n \times n}$ such that $U_{A_k} := P_k A Q_k$ and $U_{B_k} := Q_k^H B_k^* P_k^H$ are two upper triangular matrices. Since $\|P_k\| = \|Q_k\| = 1$, $\|U_{A_k}\|_F = \|A\|_F$, and $\|U_{B_k}\|_F = \|B_k\|_F \leq \|B\|_F + \frac{s-r}{k}$, it follows from the Bolzano–Weierstrass Property [12] that there exists a convergent subsequence of (Q_k, P_k, B_k) such that its limit is equal to (Q, P, B) for some unitary matrices Q and P , which completes the proof of part 1.

From part 1, there exist two unitary matrices P and Q giving rise to

$$\begin{aligned} (B^\top \otimes A)\mathcal{P} &= (Q \otimes P^H)(U_B \otimes U_A)(P \otimes Q^H)\mathcal{P} \\ &= (Q \otimes P^H)(U_A \otimes U_B)\mathcal{P}(Q^H \otimes P). \end{aligned}$$

Thus part 2 holds.

Let the diagonal entries of U_A and U_B be denoted by $\{a_{ii}\}$ and $\{b_{jj}\}$, respectively. Then, $(U_A \otimes U_B)$ is an upper triangular matrix with given diagonal entries, specified by a_{ii} and b_{jj} . Here, one can assume without loss of generality that U_A and U_B are $n \times n$ matrices, i.e., $m = n$. After multiplying $(U_A \otimes U_B)$ with \mathcal{P} from the right, the position of the entry $a_{ii}b_{jj}$ is changed to be in the $j + n(i-1)$ -th row and the $i + n(j-1)$ -th column of the matrix $(U_A \otimes U_B)\mathcal{P}$. They are then reshuffled by a sequence of permutation matrices to form a block upper triangular matrix with diagonal entries arranged in the following order

$$\left\{ a_{11}b_{11}, \begin{bmatrix} 0 & a_{11}b_{22} \\ a_{22}b_{11} & 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 & a_{11}b_{nn} \\ a_{nn}b_{11} & 0 \end{bmatrix}, a_{22}b_{22}, \begin{bmatrix} 0 & a_{22}b_{33} \\ a_{33}b_{22} & 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 & a_{nn}b_{22} \\ a_{22}b_{nn} & 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 & a_{n-1,n-1}b_{nn} \\ a_{nn}b_{n-1,n-1} & 0 \end{bmatrix}, a_{nn}b_{nn} \right\} \quad (8)$$

By (8), it can be seen that

$$\begin{aligned} \sigma((B^\top \otimes A)\mathcal{P}) &= \left\{ a_{ii}b_{ii}, \pm \sqrt{a_{ii}a_{jj}b_{ii}b_{jj}}, 1 \leq i, j \leq n \right\} \\ &= \left\{ \lambda_i, \pm \sqrt{\lambda_i \lambda_j}, 1 \leq i, j \leq n \right\} \end{aligned}$$

where $\lambda_i = a_{ii}b_{ii} \in \sigma(A^\top B)$ for $1 \leq i \leq n$. Note that if $m \neq n$, the diagonal entries given in (8) will remain unchanged, except that some zero diagonal entries will be added or removed from the diagonal.

Note that when $m = n$, the decomposition given in part 1 of Lemma 2.2 is called *the periodic Schur decomposition* [3]. The following result, providing the unique solvability conditions of (1), is an immediate consequence of Lemma 2.2.

Theorem 2.1. *The \top -Stein matrix equation (1) is uniquely solvable if and only if the following conditions are satisfied: for $\lambda \in \sigma(A^\top B)$ and $\lambda \neq -1$, $\frac{1}{\lambda} \notin \sigma(A^\top B)$; -1 can be an eigenvalue of the matrix $A^\top B$, but must be simple.*

PROOF. From (6), we know that the \top -Stein matrix equation (1) is uniquely solvable if and only if

$$1 \notin \sigma((B^\top \otimes A)\mathcal{P}). \quad (9)$$

By Lemma 2.2, if $\lambda \in \sigma(A^\top B)$, then $\frac{1}{\lambda} \notin \sigma(A^\top B)$. Otherwise, $1 = \sqrt{\lambda \cdot \frac{1}{\lambda}} \in ((B^\top \otimes A)\mathcal{P})$. On the other hand, if $-1 \in \sigma(A^\top B)$ and -1 is not a simple eigenvalue, then $1 \in \sigma((B^\top \otimes A)\mathcal{P})$. This verifies (9) and the proof of the theorem is complete.

3. Smith-type iterative methods

Originally, Smith-type iterative methods [8] are developed to solve the Stein equation

$$X = AXB + C, \quad A \in \mathbb{R}^{m \times m}, B \in \mathbb{R}^{n \times n}, C \in \mathbb{R}^{m \times n}.$$

Our main thrust in this section is to extend the Smith-type iterative methods to solve (2) by means of the formula (5) and also preserve the convergence properties embedded in the original methods. Here we summarize the iterative methods as follows.

- **The r -Smith iteration methods:**

$$X_{k+1} = \sum_{i=0}^{r-1} A_k^i X_k B_k^i,$$

where

$$\begin{aligned} X_0 &= AC^\top B + C, & A_0 &= AB^\top, & A_{k+1} &= A_k^r, & k &\geq 0, \\ B_0 &= A^\top B, & B_{k+1} &= B_k^r, & k &\geq 0. \end{aligned}$$

It can be seen that if $\rho(A^\top B) = \rho(AB^\top) < 1$, the r -Smith iteration converges to the unique solution of (1)

$$X = \sum_{i=0}^{\infty} (AB^\top)^i (AC^\top B + C) (A^\top B)^i.$$

When $r = 2$, the r -Smith iteration is also called **the Smith accelerative iteration** [7] and it has been shown in [8] that the Smith accelerative iteration performs more effective than any other r -Smith iterations. One possible drawback of the Smith-type iterative methods is that it cannot handle the case when the eigenvalue -1 of $A^\top B$ is simple. Based on the solvable conditions given in this work, it is possible to develop a specific technique working on the particular case and it is a subject currently under investigation.

4. Conclusion

In this paper, we obtain necessary and sufficient conditions for the existence and uniqueness of the solution of the \top -Stein equation and find that the solvability conditions between (1) and (5) are different. When the equation (1) is consistent, we extend the capacity of Smith-type iteration methods from the Stein equation to the \top -Stein equation.

References

- [1] H. Braden, The equations $A^\top X \pm X^\top A = B$, SIAM Journal on Matrix Analysis and Applications 20 (2) (1998) 295–302.
- [2] X. Ma, J. Han, F. Pu, A study of solution existence for matrix equation $AX + X^\top C = B$, Journal of Shenyang Institute of Aeronautical Engineering 4 (2003) 64–66.
- [3] A. W. Bojanczyk, G. H. Golub, P. Van Dooren, Periodic schur decomposition: algorithms and applications, in: Proceedings of SPIE Conference, Vol. 1770, San Diego, CA, USA, 1992, pp. 31–42.
- [4] C.-Y. Chiang, E. K.-W. Chu, W.-W. Lin, On the \star -Sylvester equation $AX \pm X^\star B^\star = C$, Appl. Math. Comput. 218 (17) (2012) 8393–8407.
- [5] Y. Su, G. Chen, Iterative methods for solving linear matrix equation and linear matrix system, Int. J. Comput. Math. 87 (4) (2010) 763–774.
- [6] M. Wang, X. Cheng, M. Wei, Iterative algorithms for solving the matrix equation $AXB + CX^\top D = E$, Appl. Math. Comput. 187 (2) (2007) 622–629.
- [7] R. A. Smith, Matrix equation $XA + BX = C$, SIAM J. Appl. Math. 16 (1968) 198–201.
- [8] B. Zhou, J. Lam, G.-R. Duan, On Smith-type iterative algorithms for the Stein matrix equation, Appl. Math. Lett. 22 (7) (2009) 1038–1044.
- [9] D. S. Bernstein, Matrix mathematics, 2nd Edition, Princeton University Press, Princeton, NJ, 2009, theory, facts, and formulas.
- [10] R. A. Horn, C. R. Johnson, Topics in matrix analysis, Cambridge University Press, Cambridge, 1994, corrected reprint of the 1991 original.
- [11] F. Zhang, Matrix Theory: Basic Results and Techniques, Universitext (1979), Springer, 1999.
- [12] W. Rudin, Principles of mathematical analysis, 3rd Edition, McGraw-Hill Book Co., New York, 1976, international Series in Pure and Applied Mathematics.